Construction of ceRNA networks in endocrine therapy resistance model systems

Saheli Pramanik [1,2*], Partha Das [1,2*], Monalisa Mukherjee [1,2], Kartiki V. Desai [1,#]

[1] Biotechnology Research Innovation Council-National Institute of Biomedical Genomics (BRIC-NIBMG), Kalyani, India

[2] Ph.D. Program, Regional Centre for Biotechnology, 3rd Milestone, Faridabad-Gurugram Expressway, Faridabad, India

**These authors have contributed equally to the manuscript*

Correspondence to
Kartiki V. Desai
National Institute of Biomedical Genomics
P.O. N.S.S.
Kalyani 741251, India
Email- kd1@nibmg.ac.in
Telephone: +91-33-29772151 Ext 264

ABSTRACT

Endocrine therapy resistance (ETR) in breast cancer is achieved by decreasing expression or acquiring mutations in ER, dysregulating cell cycle genes, and/or ER co-activators/co-repressors. We reported that high expression of JMJD6 induced ETR by depleting ER. In this study, RNA-sequencing of parental MCF7, Tamoxifen resistant (TAMR), Long-term Estrogen deprived (LTEDI), and Jumonji domaining containing protein 6 (JMJD6) overexpressing (JOE) cells was carried out. 170 differentially expressed genes common to all cells were found, of which 73 maintained the same directionality in expression. Pathway based curation identified 7 up- and 14 down-regulated genes spanning multiple cancer pathways. These genes were used for CNC and ceRNA network construction and a triad FLT4:MIR503HG:miR497/195/424 was discovered in TAMR-LTED cells, but triads were rare in JOE cells. ER expressing cells probably devise a complex ceRNA based mechanism for ETR establishment, whereas lowering ER is sufficient in JOE cells. Further, the clustering of TCGA samples based on the expression of upregulated ETR genes led to the bifurcation of ER+ tumors into two groups, one intermixing with ER- tumors and showing poorer survival. Of the 7 up genes, B3GNT3, ADORA2B, SLC2A3, and FLT4 showed high expression in ETR models and normal breast but lower expression in tumors. We suggest that re-expression of these genes occurs during therapeutic intervention leading to poorer outcome. In addition to these 4 genes, testing for ER, MIR503HG, miR497, miR195, and miR424 expression could be useful in ETR prediction.

KEYWORDS: ceRNA, Tamoxifen, LTED, JMJD6, CNC network, lncRNA

1. INTRODUCTION

Breast cancer (BrCa) is the most common cancer in women and has high mortality rates due to high cellular and molecular heterogeneity (1,2). 70% of diagnosed breast cancer is estrogen receptor positive (ER+) and is treated with endocrine therapy (ET) in the adjuvant setting. Selective Estrogen Receptor Modulators (SERMs) such as Tamoxifen (TAM), Selective Estrogen Receptor Degraders (SERDs) like Fulvestrant and Aromatase Inhibitors (AIs) Letrozole, and Anastrozole have successfully been used as first line of therapy (3). However, about 30-40% of women recur and become non-responsive to endocrine therapy (4). Current treatment options for these women are very limited and ET resistance (ETR) remains a major challenge for the treatment of both recurrent and metastatic breast cancer (MBC). Development of tests that forecast ETR and identify novel targets for treating resistant disease is a major area of research.

Several genetic tests such as Oncotype Dx, EndoPredict, and MammaPrint provide the risk of recurrence scores and many studies have developed gene signatures associated with TAM, Fulvestrant, and ETR in breast cells and cell line xenografts via transcriptome analysis (5–9). Few of these signatures have been adapted in the clinic to estimate ETR because rarely do tumors display inherent resistance. Such resistance is acquired during the course of treatment and transcriptomic profiles during treatment from repeated biopsies are unavailable. Using cell line and patient sample data, pathways leading to such resistance are well worked out and extensively described in the literature (10). Molecular heterogeneity, differential response and outcomes to the same therapy, and use of complex pathways in ETR, limit the use discovered gene signatures in the clinic. Broadly, events leading to ETR can be grouped into 3 types- 1) lowering of ER expression, 2) acquired resistance by ligand-independent activation of ER and/or high expression of cell cycle genes, or epigenetic changes, and 3) by gain of activating mutations in ER itself. Different modes of ETR involve distinct factors that elicit different gene expression programs.

We propose that an integrated RNA network analysis to find common genes/pathways using cell lines representing various forms of ETR could be useful in designing biomarkers that indicate potential ETR, independent of the initiating event. It is clear that the expression of not only the coding mRNAs but also the non-coding genome influences ETR (11). While the long non-coding RNAs (lncRNAs) regulate mRNA expression by cis-/trans- regulation at the transcriptional level, microRNAs (miRNAs) work post-transcriptionally to interfere with the translation of the protein encoded by its target mRNA (12). Often the same miRNA has sequences complementary to both lncRNAs and mRNAs. Various regulatory outcomes are possible based on the levels of miRNA, lncRNA, and mRNAs present at any given time (13). Clearly, the ratio of mRNA:lncRNA:miRNA will tend to determine the final expression levels of a gene/protein. Typically, co-expression and competing endogenous RNA networks are used to understand the complex RNA regulation and they have been successfully used to segregate ER+ and ER- tumors, understand the response to Tamoxifen in cancer patients and generate lncRNA or miRNA signatures (11,14–17). ceRNA networks for Tam sensitive and TAMR

cells have been constructed previously and MCF7 derived LLC3 and LCC9 Tam resistant cells (18,19). ceRNA networks for alternate modes of ETR have not been constructed earlier.

Here we generate total transcriptomes of ETR cell line models. MCF7 derived Tamoxifen resistant (TAMR), Long-Term Estrogen Deprived-Independent (LTEDI) that do not require estrogen for proliferation (20), and JOE cells that overexpress Jumonji Domain Containing protein 6 (JMJD6). Our laboratory has developed TAMR cells that have high ER, RET, and phospho-AKT levels, LTEDI cells having high ER, ERK1/ERK2 expression whereas JOE cells harboring low ER, high RET and high ERK1/ERK2 levels. JOE cells are insensitive to endocrine therapy and re-sensitized is possible by depleting JMJD6 (21). Incidentally, TAMR and LTED cells also showed high expression of JMJD6 and depletion of JMJD6 in TAMR cells restored sensitivity to TAM (21). JMJD6 is a major epigenetic regulator of gene expression, and induced cyclin E2, E2F targets, and G2/M checkpoint genes to promote cell cycle advancement in breast cancer cells, thus removing their dependency on ER (21–23). Increased cyclin E2 was also identified as a marker of prognosis and persistent Fulvestrant resistance in MBC (24). Clearly, each cell model system can evade TAM induced growth restriction via overlapping and distinct modes of ETR. In this study, genes common and similarly regulated in all 3 cell lines were identified, filtered by pathway-based curation, and subjected to CNC and ceRNA network analysis.

2. MATERIALS AND METHODS

2.1. Cell culture

MCF7 cells were purchased from the American Type Culture Collection (ATCC, VA, USA). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (GIBCO, USA) with 5% fetal bovine serum (FBS) (GIBCO, USA) and 1% PenStrep (GIBCO, USA) in humidified 5% CO_2 incubator at 37°C. Cell lines were negative for mycoplasma presence (Lonza, Switzerland). Generation of JMJD6 overexpression (JOE) cells, TAMR, and LTEDI cells are described previously (21).

2.2. Cell Viability and Colony Forming Assay

Cell viability assay was performed by CCK8 kit (Sigma Aldrich, USA) according to the manufacturer's protocol. MCF7 cells were starved for 3 days in estrogen-deprived media. After 3 days, MCF7 and LTEDI cells were plated at a density of 10^4 cells/well in 96 well plate. Treatments were 1μM 4-hydroxy TAM (4-OHT) and ethanol (as vehicle control) for 72 hours. Absorbance was measured at 450 nm**.** LTEDI cells (10^3 cells/well) were seeded in 6-well plates. The next day, cells were treated with ethanol/1 μM 4-OHT for 9 days with a daily media change. Cells were fixed with 4% paraformaldehyde and stained using a crystal violet solution. The number of colonies was quantified and cells were photographed.

2.3. RNA-sequencing, Library Preparation and Data Analysis

RNA was isolated from MCF7, TAMR, and LTEDI cells by using a Qiagen RNAeasy mini kit (Qiagen, Germany). Library preparation was carried out at the NIBMG Core facility for 100 bp paired-end reads (75 million each) using NOVASeq. Following FASTQC quality check FAST Q files were trimmed with CutAdapt to remove the adapter sequences and aligned with the human genome (GrCh38.1) by Hisat2.0 algorithm. Alignments were processed by the FeatureCounts package for obtaining RNA/ transcript counts for differentially expressed coding (DEGs) and for non-coding RNAs (DELs) separately. Data is available at GEO Omnibus (GSE255704). Limma-voom analysis followed by Benjamini-Hochberg correction was used to obtain DEGs/DELs significant at adjusted p-value ≤ 0.05 and log fold change= 1(log2FC) (https://usegalaxy.org). Cluster 3.0 and Tree View programs were used for heatmap generation and visualization. Pathway enrichment analysis was done by EnrichR (https://maayanlab.cloud/Enrichr/). Top 10 significant pathways are shown for further comparison. Chord diagrams for genes appearing in various pathways was constructed using SRplot. To obtain DEG data from Fulvestrant resistant MCF7 cells, GSE118713 was analysed using GEO2R and our internal pipeline (25).

2.4. Co-expression (CNC) and ceRNA networks

To associate the lncRNAs with target mRNAs, we constructed the lncRNA-mRNA co-expression network. For each pair of DEGs and DELs, the Pearson correlation was calculated and those pairs with correlations of 0.98 or greater were used to construct the network. For cis-regulation, the genomic distance between each pair was calculated using data obtained from Biomart (https://www.ensembl.org/biomart/martview/). None of the pairs were located within 10 kb (data not shown). For the ceRNA networks (mRNA-lncRNA-miRNA), first, the miRNAs for DELs were predicted from the miRcode (http://www.mircode.org) and both miRcode and TargetScan (v7.2) (https://www.targetscan.org) database were used for miRNAs and DEGs. A

list of mRNA-lncRNA-miRNAs common to both TAMR-LTED was obtained. JOE networks were constructed separately. Pathway based curation was carried out to obtain DEGs that appeared in multiple significant pathways to give 7 upregulated and 14 downregulated genes (21 genes). The mRNA-lncRNA-miRNA interaction networks were constructed using 21 DEGs and corresponding lncRNAs and miRNAs. Cytoscape (https://cytoscape.org, version 3.10.0) was used to construct and visualize all networks.

2.5.TCGA Data and Survival Analysis

cBioPortal (https://www.cbioportal.org/) was used to generate the gene expression heatmap for all breast cancer subtypes using The Pan-Cancer Atlas dataset. Clinical data, age, subtype information, inclusion-exclusion criteria etc of 1084 samples are available at https://www.cbioportal.org/study/clinicalData?id=brca_tcga_pan_can_atlas_2018. mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM) were used to determine gene expression patterns of select genes. The relative expression of selected genes as tumor versus normal was plotted by UALCAN (https://ualcan.path.uab.edu). The median expression value was used as a cut-off and significant difference between groups was assessed using the unpaired two-tailed Student's t-test. The prognostic impact of selected genes was analysed using the Kaplan-Meier Plotter (http://kmplot.com/analysis/) using ER+ breast cancer treatment naïve datasets.

3. RESULTS

3.1.Effect of Tamoxifen on LTEDI cells

Both agonistic and antagonistic activity of TAM has been reported for LTEDI cells in the literature, however, independence from the requirement of exogenous E_2 for proliferation is common to all LTEDI cells (19). To determine the nature of our cells, we studied the effect of 4-OHT both in short-term cultures using CCK-8 assays and in longer-term cultures by colony forming assay. While parental MCF7 cells desist, LTEDI cells continue to proliferate in the presence of 4-OHT over 72 hours (Figure 1A). Secondly, the overall proliferation rate of LTEDI is higher than parental MCF7 cells in the absence of 4-OHT, and this is further enhanced by the addition of the drug (Figure 1B). Our data indicates that 4-OHT may serve as an agonist in LTEDI cells.

3.2. Identification of commonly regulated DEGs in JOE, TAMR and LTEDI cells

The workflow for identifying mRNA, lncRNA, and miRNA of our analysis is shown in figure 2. We published the transcriptomic program of JOE cells expressing low ER earlier and identified about 1100 genes that were differentially expressed in these cells (21). RNA-seq analysis of TAMR cells identified 8685 genes as significantly differentially expressed (adjusted p-value ≤ 0.05) when compared to parental MCF7 cells. 1609 genes were upregulated and 2332 genes were down-regulated ($\log_2$FC=1) (Figure 3 A, B supplementary file 1). In LTEDI cells, 5630 DEGs were found (adjusted p-value ≤ 0.05), of which, 687 genes were upregulated and 1476 genes were down-regulated (log2FC=1) (Figure 3 A, B supplementary file 1). Pathway enrichment analysis (EnrichR) was used to compare the two cell line models and significant pathways are shown (Adjusted p-value ≤ 0.05, Figure 3 C). Estrogen response early, Estrogen response late, and Interferon alpha and gamma pathways were regulated in both TAMR and LTEDI cells, however, the number of genes was variable. These are indicated by the size and color variation in the bubble plot (Figure 3 C). mTORC1 signalling, p53 pathway, Hypoxia, Glycolysis, Cholesterol homeostasis, Xenobiotic metabolism, Epithelial-Mesenchymal Transition (EMT) were unique to TAMR cells while TNF-alpha signalling via NFkB pathways, KRAS signalling Up, and Wnt-beta Catenin signalling was found in LTEDI cells (Figure 3 C).

Our goal was to obtain genes that accounted for an escape from drug induced growth restriction. The intersection of TAMR and LTEDI DEGs yielded 1482 genes. Interestingly 94% of these genes showed the same directionality in regulation; with 453 induced and 973 suppressed DEGs. TAMR-LTEDI-JOE intersection had 170 genes common amongst them, with 73 genes sharing the directionality (all up/all down/all cells) and referred to as the endocrine therapy resistance cassette genes (ETR cassette) (Figure 3 D). Out of 73 genes, 33 were upregulated and 40 downregulated (Figure 3 E). The remaining 97 genes out of 170 genes were regulated in the opposite direction in JOE cells and mostly contained the ER regulated genes. In JOE, ER levels are suppressed but ER is robustly expressed in both TAMR and LTEDI cells, which may partially account for this observation. Since Fulvestrant is used for endocrine therapy, we carried out a similar analysis (both by GEO2R and our pipeline) in Fulvestrant resistant MCF7 cell data described in GSE118713 (25). After comparison with TAMR-LTEDI-JOE data, we found 85 genes were common (Supplementary figure 1 A and B), 26 upregulated and 11 downregulated genes.

The ETR cassette was enriched for KRAS signalling, Hypoxia, Glycolysis, Inflammatory response pathways in the upregulated genes and Myogenesis, TNF-alpha signalling via NFkB, Estrogen response late, Interferon gamma response represented in the downregulated genes.

We noticed that 7 induced DEGs and 14 downregulated DEGs participated in every significant pathway identified, indicating their importance in ETR (Figure 3 E). Next, we assayed the expression of the ETR cassette in TCGA data and used pathway-based curated genes (7 up:14 down) to construct CNC and ceRNA networks.

3.3. Identification of Differentially Expressed lncRNAs (DELs)

For the construction of CNC networks, we used the 7 up and 14 downregulated genes since they appeared in every significant pathway identified from DEGs and most were annotated as cell surface receptor proteins. If they were involved in ETR, they could be exploited as novel therapeutic targets. To construct CNC networks, we first estimated the expression of long non-coding RNAs in TAMR, LTEDI, and JOE cells (Figure 4 A, B). In JOE cells very few DELs were obtained (85 upregulated and 7 down-regulated) and very few of the lncRNAs overlapped with either TAMR or the LTEDI DELs. This suggests that JMJD6 may modulate the ETR cassette genes directly or by an alternate lncRNA mechanism distinct from TAMR-LTEDI.

959 DELs were identified in TAMR cells of which, 448 lncRNAs were upregulated and 511 lncRNAs were downregulated (adjusted p-value $\leq$ 0.05, $\log_2$FC=1) (Figure 4 A, B supplementary file 1). In LTEDI cells, 527 lncRNAs were differentially expressed, of which 198 were up and 329 were downregulated (adjusted p-value $\leq$ 0.05, $\log_2$FC=1) (Figure 4 A, B, supplementary file 1). Among DELs, 337 were common between TAMR and LTEDI cells and like the observation in DEGs, maintained the common directionality. 129 were upregulated and 200 were downregulated (Figure 4 B).

3.4.Construction of co-expression network

LncRNAs affect mRNA expression in both cis- and trans-regulatory manner. Positive correlation typically indicates cis-regulation of the mRNA by the lncRNA if the pairs are localized in close proximity to the genome (within 10 kb). To find such pairs, analysis of DEG:DEL pairs where both were upregulated, or both were downregulated was carried out by using Pearson correlation analysis, and positively correlated pairs were studied (Pearson correlation cut off > 0.98 for TAMR-LTEDI and > 0.95 for JOE). Four different DEG and DEL co-expression networks were constructed using the selected 7 up and 14 down DEGs in TAMR and LTEDI cells. 7 up DEGs positively correlated with 190 and 60 DELs in TAMR and LTEDI cells respectively (Supplementary figure 2 A, B). 14 downregulated DEGs had a positive correlation with 239 DELs in TAMR and 177 DELs in LTEDI (supplementary figure 2 C, D). As with DEGs, we were more interested in common rather than cell line-specific networks and identified a set of DELs that are commonly co-expressed in both TAMR and LTEDI. Specifically, 25 up-regulated DELs were co-expressed with 7 up DEGs (Figure 5 A), and 14 downregulated genes correlated with 93 common DELs (Figure 5 B). In JOE cells, only 5 DELs were common and only 3/7 upregulated genes had 55 positively correlated DELs, and 5/14 downregulated DEGs paired with 4 DELs (Figure 5 C, D). To establish cis-regulation, the genomic distance between the DEG:DEL pairs was calculated. None of the positively correlated mRNA-lncRNA pairs were placed near each other (within 10 kb), which suggested that the lncRNAs did not regulate the mRNAs via cis-regulation. Therefore, it is plausible that in ET resistance, cis-regulatory mechanisms play less of a role.

3.5.Construction of competing endogenous RNA (ceRNA) network

Since lncRNAs act as an endogenous "sponge" to sequester microRNAs (miRNAs), thereby altering mRNA expression, we constructed ceRNA networks to explore the mRNA:lncRNA:miRNA connectivity. 21 DEGs (7 upregulated and 14 downregulated) from pathway-based curation were used. Since we did not have data for miRNA profiles for these cells, we used the following conditions to predict miRNA levels from the positively correlated DEG:DEL pairs-

1) if both lncRNA and mRNA from the CNC network are upregulated, it is likely that the miRNAs regulating them are depleted.
2) if both are low, miRNAs inhibiting them are likely upregulated.

Since TAMR-LTEDI cells showed a great degree of overlap but JOE had different DEG:DEL pairs, we made separate ceRNA networks. The TAMR/LTEDI/up network consisted of 15 miRNAs, 5 DEGs, and 5 DELs whereas, 11 miRNAs, 6 DEGs, and 8 DELs were found in the TAMR/LTEDI/down network (Figure 6 A and B respectively, Table 1). Together, 15 miRNAs are predicted to be downregulated and 11 miRNAs upregulated.

Next, we identified negatively correlated DEG-DEL pairs (cut off > - 0.98). For negatively correlating DEG:DEL pairs,

1) miRNA binding could cause degradation of interacting lncRNAs or interacting mRNAs (26,27).
2) miRNAs binding to mRNAs may facilitate miRNA-driven upregulation of lncRNAs and vice-versa.
3) The negatively correlated network may represent a set of shared miRNAs that play a role in regulating both mRNAs and lncRNAs.

Figure 6 C shows the network of 6 upregulated DEGS in both TAMR and LTEDI paired with 9 downregulated DELs and 11 corresponding miRNAs. For 8 Downregulated DEGs common in both TAMR and LTEDI, 4 upregulated DELs and 3 miRNAs were obtained (Figure 6 D). From these studies, 14 miRNAs are predicted to be upregulated. In JOE cells, a separate ceRNA network was constructed using upregulated pairs, consisting of 12 miRNAs, 3 DEGs, and 3 DELs (Figure 6 E). We did not find any overlapping miRNAs for downregulated DEG-DEL pairs of JOE sets.

3.6.Validation of the miRNA predictions

Since we did not have the miRNA expression data, we used ceRNA networks to predict the most likely expression level of the 40 interacting miRNAs in these cells. We predicted 25 to be upregulated and 15 to be downregulated. To validate our observations, we used publicly available miRNA expression data from GSE 56411 representing the regulation of miRNA in 3 different TAMR clones (28). Data from JOE and LTEDI cells was not available. 24/40 miRNAs were represented in the TAMR clone dataset and their ceRNA networks are shown in Supplementary figure 3. In 2 out of 3 TAMR clones, 7/11 miRNAs predicted to be upregulated and 6/13 predicted to be downregulated by our ceRNA network analysis showed

the same directionality in expression (Figure 7). In JOE cells, 2/12 downregulated miRNAs overlapped with the validated TAMR-LTEDI miRNAs. These data underline the importance of constructing ceRNA networks and the resultant predictions. Table 2 summarizes the mRNA-lncRNA-miRNA relationships for the validated miRNAs.

3.7.Gene expression in breast cancer patient samples

In the publicly available pan cancer (TCGA breast) dataset, we analysed the Dead/Alive clinical status and ER+ subtype specific gene expression data using death as an outcome to represent a poor response to treatment. We failed to obtain DEGs, but this was expected since ET resistance is acquired over prolonged exposure to drugs and long-term estrogen deprivation, and it is rarely intrinsic (data not shown). Since very little data comparing longitudinally collected samples at diagnosis and following recurrence is available, we tested if our ETR cassette was better at discriminating the ER+ tumors at diagnosis. Since this set originated from a cell model that had an E_2-ER independent axis of growth, we hypothesized that some ER+ tumors with the potential to avoid ET would cluster better with ER- tumors. Supervised hierarchical clustering analysis of 1084 samples (all subtypes) using mRNA expression z-scores relative to all samples (log RNA Seq. V2 RSEM) for 73-gene ETR cassette was carried out and is shown in Figure 8 A (supplementary figure 4 A). Interestingly, the 33 upregulated ETR split the ER+ subtype into two groups, one in which ER+ tumors were mixed with TNBCs, and another that remained distinct from ER- tumors. However, the down-regulated genes showed no consistent pattern of expression (Supplementary figure 3 A). This indicates that this 33 gene ETR set obtained from cellular models of ET resistance may be useful in the segregation of ER+ tumors into those that respond versus those that do not. However, more studies are required to validate this observation.

We checked the expression of these 33 upregulated genes in the UALCAN database and the association of these genes with patient survival using KM plots (Figure 8 B, C, and supplementary figure 4 B, C). As compared to normal, CXCR4, RGS16, and LCP1 were highly expressed in tumors but this indicated better overall patient survival. In contrast, B3GNT3, ADORA2B, SLC2A3, and FLT4 were less expressed in tumors when compared to normal samples but their high expression was associated with inferior overall survival. However, these genes were highly expressed in cells that showed ETR and in a sub-set of ER+ tumors from TCGA. It is tempting to suggest that increased re-expression of these 4 genes in tumor tissues during treatment-induced progression may indicate their escape from ET responsiveness. Testing the expression of these genes may be of value to determine an unfavorable prognosis during ET.

4. DISCUSSION

Cells achieve ETR by different modes and using TAMR/Fulvestrant resistant cells and ETR xenograft models, gene signatures have been developed but have not been deployed extensively in the clinical setting. Information from other modes of ETR, such as lowering of ER expression or gain of agonistic activity in resistant cells is lacking. We hypothesized that using combined RNA regulatory networks from cells representing distinct mechanisms to resistance will improve the clinical adaptability of such gene signatures. We established and used three model systems TAMR, LTEDI, and JOE for transcriptomic analysis and subsequent CNC and ceRNA network construction. Of these, TAMR and LTEDI cells are the most used endocrine therapy model systems (29). In LTEDI cells, the Tam effect is controversial since both agonistic and antagonistic activity has been reported. In our cells, we report that Tam behaves as a pure agonist. This allowed us to integrate 3 different modes of ETR.

To find commonly regulated genes, we analyzed RNA-Seq data of TAMR, LTEDI, and JOE cells and identified 170 DEGs and 73 ETR cassette genes were expressed in the same direction in all 3 cell lines. Pathway enrichment analysis identified crucial cancer hallmark pathways. For CNC and ceRNA analysis, we identified genes that were represented in multiple pathways and such pathway-based curation resulted in 7 common upregulated and 14 common downregulated genes. Most genes were cell surface receptors and involved in various types of cancers (Table 3). CXCR4, B3GNT3, and ADORA2B are poor prognostic markers in breast cancer and therefore good targets for cancer treatment (30–32). Though Fulvestrant resistant cells use increased EGFR and other tyrosine kinase member-driven ETR, we compared this data to ours and found 85/170 genes in GSE118713 profiles (33,34). 26 genes out of 85 were expressed in the same direction.

We first constructed a CNC network using DEGs and DELs in both TAMR-LTEDI cells and separately in JOE cells. We failed to find cis-regulation of mRNAs by lncRNAs as a mechanism for co-expression and suggest that this may not be a preferred regulatory mechanism in ETR. The most relevant ceRNA networks were obtained by pathway-based curation of the correlated DEG:DEL lists. This led to the selection of 7 induced and 14 suppressed genes. Generating ceRNA networks using both positively and negatively correlated DEG:DEL pairs, we obtained a total of 27 lncRNAs. We did not have miRNA expression data but based on the ceRNA hypothesis, we predicted miRNA expression based on the levels of corresponding mRNAs and lncRNAs. We obtained a total of 40 miRNAs in this network and found 24 of them in TAMR miRNA data (28). Of these, we could correctly predict the expression level of 7/11 upregulated and 6/13 downregulated miRNAs. Most miRNAs identified in this study have been previously described in ETR or implicated in breast cancer prognosis (35). These miRNAs can be quantified in both normal healthy individuals as well as in miRNA-based signatures in cancer patients via liquid biopsy approaches (Table 2). In JOE cells, we obtained very few DELs. We propose that JMJD6 uses the depletion of ER as the major ETR mechanism rather than altering ceRNA networks. TAMR and LTEDI cells on the other hand have abundant ER, which may provoke more complex ETR mechanisms. For example, Vilquin *et al* showed that depletion of miR424 and upregulation of miR205 resulted in Aromatase Inhibitor resistance phenotype in ER+ cells (36). We predicted downregulation of miR195/miR497. These miRNAs are embedded in lncRNA MIR497HG and their downregulation was associated with TAM resistance (37). Both miRNAs contain ER binding sites in their promoter and they are upregulated by E_2 treatment. In Tam presence, their expression would be abrogated. Curiously, none of the ETR cassette genes have been described previously in TAMR signature genes (8,38,39), nor did they overlap with the EndoPredict (3

proliferation and 5 hormone receptor-related genes from Myriad Genetics), nor with the ENDORSE prognostic model for ET that identified MYC targets, E2F1 targets, and PI3KC-mTOR signalling pathway gene expression in tumors as high risk for ETR (40,41). Our network suggests that these pathways from ENDORSE could be represented in an alternate manner. miR497-5p and miR195 interfere with PI3K-mTOR signalling. We predicted their lower expression in resistant cells based on increased expression of DEL MIR503HG in TAMR-LTEDI cells. MIR503HG is known to sponge the interfering miRNAs allowing PI3K-mTOR activation and subsequent ETR (42). Moreover, miR424 arises from MIR503HG and its reduced expression as predicted by our network allows increased expression of VEGF that participates in neo-angiogenesis of tumors (42,43). FLT4 is the receptor for VEGF and it is one of the 7 ceRNA network genes that displayed high expression in resistant cells. Together, the network of FLT4, MIR503HG, miR424, and miR497/195 indicate the abolishment of E_2 activity, augmentation PI3K-mTOR signalling, and high FLT4-VEGF axis for possible angiogenesis during ETR. Together, these data significantly validate the usefulness of the ceRNA network-based approach to predict actionable miRNAs.

We next checked the status of the ETR cassette genes in TCGA datasets. We hypothesized that if ER+ tumors were to be resistant to ET, they should display certain gene expression profiles like ER- tumors as they no longer use the E2/ER axis for sustenance. When we assessed the expression of ETR cassette genes in TCGA breast cancer data, the expression pattern of upregulated genes led to the clustering of a subset of ER+ tumors with ER-negative breast cancers (Figure 8). This suggests that genes appearing in ETR cell data may be able to segregate the ER+ tumors. Next, to determine the individual profile of each gene, we employed the UALCAN database. CXCR4, RGS16, and LCP1 displayed higher expression in tumors when compared to normal breast but did not associate with poor outcomes in ER+ breast cancer. Interestingly, B3GNT3, ADORA2B, SLC2A3, and FLT4 expression was lower in tumors and higher in normal breast, and low expression was associated with better survival outcomes. We posit that their re-expression of these genes in the tumor may indicate potential ETR. Further studies using immunohistochemistry or RT-PCR analysis can test this hypothesis.

5. FUNDING

This work was supported by Intramural Funding (Grant 60102) from National Institute of Biomedical Genomics. Saheli Pramanik is a graduate student supported by a UGC fellowship (211610013835). Partha Das is a graduate student supported by a CSIR fellowship (09/1033(0010)/2019-EMR-I). Monalisa Mukherjee is a graduate student supported by a DBT fellowship (DBT/2021-22/NIBMG/1845).

6. AUTHOR CONTRIBUTIONS

Saheli Pramanik: Methodology, Data curation, Visualization **Partha Das**: Investigation, Visualization, **Monalisa Mukherjee**: Visualization, Investigation. **Kartiki V Desai** Conceptualization, Analysis, Supervision. **All authors** contributed to Writing- Original draft preparation, reviewing, and editing the manuscript.

7. CONFLICT OF INTEREST

The Authors declare that there are no competing interests associated with the manuscript.

8. ACKNOWLEDGEMENT

The authors would like to thank Aritra Gupta and Siddarth Bharadwaj for their technical help and assistance in editing the manuscript.

9. ABBREVIATIONS

AIs: Aromatase Inhibitors, ET: Endocrine Therapy, SERDs: Selective Estrogen Receptor Degraders, SERMs: Selective Estrogen Receptor Modulators, TAM: Tamoxifen, MBC: Metastatic Breast Cancer, JMJD6: Jumonji Domain Containing 6, JOE: JMJD6 overexpression, TAMR: Tamoxifen resistant, LTEDI: Long-Term Estrogen Deprived-Independent, RNA-seq: RNA sequencing, ceRNAs: competing endogenous RNAs, lncRNA: Long non-coding RNA, miRNA: microRNA, 4-OHT: 4-hydroxy TAM, DEGs: Differentially Expressed Coding Genes, DELs: Differentially Expressed non-coding RNAs.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.

2. Savci-Heijink CD, Halfwerk H, Hooijer GKJ, Horlings HM, Wesseling J, van de Vijver MJ. Retrospective analysis of metastatic behaviour of breast cancer subtypes. Breast Cancer Res Treat. 2015 Apr;150(3):547–57.

3. Shagufta, Ahmad I. Tamoxifen a pioneering drug: An update on the therapeutic potential of tamoxifen derivatives. Eur J Med Chem. 2018 Jan;143:515–31.

4. Zhang MH, Man HT, Zhao XD, Dong N, Ma SL. Estrogen receptor-positive breast cancer molecular signatures and therapeutic potentials (Review). Biomed reports. 2014 Jan;2(1):41–52.

5. McVeigh TP, Hughes LM, Miller N, Sheehan M, Keane M, Sweeney KJ, et al. The impact of Oncotype DX testing on breast cancer management and chemotherapy prescribing patterns in a tertiary referral centre. Eur J Cancer. 2014 Nov;50(16):2763–70.

6. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002 Dec;347(25):1999–2009.

7. Göker M, Denys H, van de Vijver K, Braems G. Genomic assays for lobular breast carcinoma. J Clin Transl Res. 2022 Dec;8(6):523–31.

8. Chanrion M, Negre V, Fontaine H, Salvetat N, Bibeau F, Mac Grogan G, et al. A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. Clin cancer Res an Off J Am Assoc Cancer Res. 2008 Mar;14(6):1744–52.

9. Saponaro M, Annunziata L, Turla A, Viganò I, De Laurentiis M, Giuliano M, et al. Extended Adjuvant Endocrine Treatment in Luminal Breast Cancers in the Era of Genomic Tests. Int J Mol Sci. 2022 Nov;23(21).

10. Haque MM, Desai K V. Pathways to Endocrine Therapy Resistance in Breast Cancer. Frontiers in Endocrinology. 2019.

11. Treeck O, Haerteis S, Ortmann O. Non-Coding RNAs Modulating Estrogen Signaling and Response to Endocrine Therapy in Breast Cancer. Cancers (Basel). 2023 Mar;15(6).

12. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature. 2012 Sep;489(7414):101–8.

13. Schwarzenbach H, Gahan PB. Interplay between LncRNAs and microRNAs in Breast Cancer. Int J Mol Sci. 2023 Apr;24(9).

14. Xiao B, Zhang W, Chen L, Hang J, Wang L, Zhang R, et al. Analysis of the miRNA-mRNA-lncRNA network in human estrogen receptor-positive and estrogen receptor-negative breast cancer based on TCGA data. Gene. 2018 Jun;658:28–35.

15. Gao C, Li H, Zhuang J, Zhang H, Wang K, Yang J, et al. The construction and analysis of ceRNA networks in invasive breast cancer: a study based on The Cancer Genome Atlas. Cancer Manag Res. 2019;11:1–11.

16. Shi W, Hu D, Lin S, Zhuo R. Five-mRNA Signature for the Prognosis of Breast Cancer Based on the ceRNA Network. Biomed Res Int. 2020;2020:9081852.

17. Venkatesh J, Wasson M-CD, Brown JM, Fernando W, Marcato P. LncRNA-miRNA axes in breast cancer: Novel points of interaction for strategic attack. Cancer Lett. 2021 Jul;509:81–8.

18. Qiao Z, Xing Y, Zhang Q, Tang Y, Feng R, Pang W. Tamoxifen resistance-related ceRNA network for breast cancer. Front cell Dev Biol. 2022;10:1023079.

19. Gao L, Shen K, Yin N, Jiang M. Comprehensive Transcriptomic Analysis Reveals Dysregulated Competing Endogenous RNA Network in Endocrine Resistant Breast Cancer Cells. Front Oncol. 2020;10:600487.

20. Masamura S, Santner SJ, Heitjan DF, Santen RJ. Estrogen deprivation causes estradiol hypersensitivity in human breast cancer cells. J Clin Endocrinol Metab. 1995 Oct;80(10):2918–25.

21. Das P, Gupta A, Desai K V. JMJD6 orchestrates a transcriptional program in favor of endocrine resistance in ER+ breast cancer cells. Front Endocrinol (Lausanne). 2022;13:1028616.

22. Lee YF, Miller LD, Chan XB, Black MA, Pang B, Ong CW, et al. JMJD6 is a driver of cellular proliferation and motility and a marker of poor prognosis in breast cancer. Breast Cancer Res. 2012;

23. Biswas A, Mukherjee G, Kondaiah P, Desai K V. Both EZH2 and JMJD6 regulate cell cycle genes in breast cancer. BMC Cancer. 2020 Nov;20(1):1159.

24. Kaminska K, Akrap N, Staaf J, Alves CL, Ehinger A, Ebbesson A, et al. Distinct mechanisms of resistance to fulvestrant treatment dictate level of ER independence and selective response to CDK inhibitors in metastatic breast cancer. Breast Cancer Res. 2021 Feb;23(1):26.

25. Achinger-Kawecka J, Valdes-Mora F, Luu P-L, Giles KA, Caldon CE, Qu W, et al. Epigenetic reprogramming at estrogen-receptor binding sites alters 3D chromatin landscape in endocrine-resistant breast cancer. Nat Commun. 2020 Jan;11(1):320.

26. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011 Sep;43(6):904–14.

27. Shetty A, Venkatesh T, Kabbekodu SP, Tsutsumi R, Suresh PS. LncRNA-miRNA-mRNA regulatory axes in endometrial cancer: a comprehensive overview. Arch Gynecol Obstet. 2022 Nov;306(5):1431–47.

28. Joshi T, Elias D, Stenvang J, Alves CL, Teng F, Lyng MB, et al. Integrative analysis of miRNA and gene expression reveals regulatory networks in tamoxifen-resistant breast cancer. Oncotarget. 2016 Aug;7(35):57239–53.

29. Chan CMW, Martin L-A, Johnston SRD, Ali S, Dowsett M. Molecular changes associated with the acquisition of oestrogen hypersensitivity in MCF-7 breast cancer cells on long-term oestrogen deprivation. J Steroid Biochem Mol Biol. 2002

Aug;81(4–5):333–41.

30. Parker CC, Kim RH, Li BDL, Chu QD. The chemokine receptor CXCR4 as a novel independent prognostic marker for node-positive breast cancer patients. J Surg Oncol. 2012 Sep;106(4):393–8.

31. Wang X, Wang Y, Chen X, He Y, Zhou X, Jiao S, et al. Identification of glycogene-based prognostic signature and validation of B3GNT7 as a potential biomarker and therapeutic target in breast cancer. J Cancer Res Clin Oncol. 2023 Dec;149(19):16957–69.

32. Sui Y, Liu J, Zhang J, Zheng Z, Wang Z, Jia Z, et al. Expression and Gene Regulation Network of Adenosine Receptor A2B in Lung Adenocarcinoma: A Potential Diagnostic and Prognostic Biomarker. Front Mol Biosci. 2021;8:663011.

33. Brünner N, Boysen B, Jirus S, Skaar TC, Holst-Hansen C, Lippman J, et al. MCF7/LCC9: an antiestrogen-resistant MCF-7 variant in which acquired resistance to the steroidal antiestrogen ICI 182,780 confers an early cross-resistance to the nonsteroidal antiestrogen tamoxifen. Cancer Res. 1997 Aug;57(16):3486–93.

34. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, et al. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. Cancer Cell. 2018 Sep;34(3):427-438.e6.

35. Howard EW, Yang X. microRNA Regulation in Estrogen Receptor-Positive Breast Cancer and Endocrine Therapy. Biol Proced Online. 2018;20:17.

36. Vilquin P, Donini CF, Villedieu M, Grisard E, Corbo L, Bachelot T, et al. MicroRNA-125b upregulation confers aromatase inhibitor resistance and is a novel marker of poor prognosis in breast cancer. Breast Cancer Res. 2015 Jan;17(1):13.

37. Tian Y, Chen Z-H, Wu P, Zhang D, Ma Y, Liu X-F, et al. MIR497HG-Derived miR-195 and miR-497 Mediate Tamoxifen Resistance via PI3K/AKT Signaling in Breast Cancer. Adv Sci (Weinheim, Baden-Wurttemberg, Ger. 2023 Apr;10(12):e2204819.

38. Hultsch S, Kankainen M, Paavolainen L, Kovanen R-M, Ikonen E, Kangaspeska S, et al. Association of tamoxifen resistance and lipid reprogramming in breast cancer. BMC Cancer. 2018 Aug;18(1):850.

39. Asghari A, Wall K, Gill M, Vecchio N Del, Allahbakhsh F, Wu J, et al. A novel group of genes that cause endocrine resistance in breast cancer identified by dynamic gene expression analysis. Oncotarget. 2022;13:600–13.

40. Nath A, Cohen AL, Bild AH. ENDORSE: a prognostic model for endocrine therapy in estrogen-receptor-positive breast cancers. Mol Syst Biol. 2022 Jun;18(6):e10558.

41. Andre F, Ismaila N, Allison KH, Barlow WE, Collyar DE, Damodaran S, et al. Biomarkers for Adjuvant Endocrine and Chemotherapy in Early-Stage Breast Cancer: ASCO Guideline Update. J Clin Oncol Off J Am Soc Clin Oncol. 2022 Jun;40(16):1816–37.

42. Han X, Li B, Zhang S. MIR503HG: A potential diagnostic and therapeutic target in human diseases. Biomed Pharmacother. 2023 Apr;160:114314.

43. Monteiro JP, Rodor J, Caudrillier A, Scanlon JP, Spiroski A-M, Dudnakova T, et al. MIR503HG Loss Promotes Endothelial-to-Mesenchymal Transition in Vascular

Disease. Circ Res. 2021 Apr;128(8):1173–90.

44. Zhang H, Li B, Zhao H, Chang J. The expression and clinical significance of serum miR-205 for breast cancer and its role in detection of human cancers. Int J Clin Exp Med. 2015;8(2):3034–43.

45. Thakur S, Grover RK, Gupta S, Yadav AK, Das BC. Identification of Specific miRNA Signature in Paired Sera and Tissue Samples of Indian Women with Triple Negative Breast Cancer. PLoS One. 2016;11(7):e0158946.

46. Zhao F, Dou Y, Wang X, Han D, Lv Z, Ge S, et al. Serum microRNA-195 is down-regulated in breast cancer: a potential marker for the diagnosis of breast cancer. Mol Biol Rep. 2014 Sep;41(9):5913–22.

47. Wang K, Yuan Y, Cho J-H, McClarty S, Baxter D, Galas DJ. Comparing the MicroRNA spectrum between serum and plasma. PLoS One. 2012;7(7):e41561.

48. Mukherjee D, Zhao J. The Role of chemokine receptor CXCR4 in breast cancer metastasis. Am J Cancer Res. 2013;3(1):46–57.

49. Zhuang H, Zhou Z, Zhang Z, Chen X, Ma Z, Huang S, et al. B3GNT3 overexpression promotes tumor progression and inhibits infiltration of CD8(+) T cells in pancreatic cancer. Aging (Albany NY). 2020 Dec;13(2):2310–29.

50. Li C-W, Lim S-O, Chung EM, Kim Y-S, Park AH, Yao J, et al. Eradication of Triple-Negative Breast Cancer Cells by Targeting Glycosylated PD-L1. Cancer Cell. 2018 Feb;33(2):187-201.e10.

51. Zhang S, Yu H, Zhang L. Role of vascular endothelial growth factor receptor-3/Flt-4 in early-stage cervical cancer. Oncol Lett. 2010 May;1(3):453–6.

52. Xiao X, Liu Z, Wang R, Wang J, Zhang S, Cai X, et al. Genistein suppresses FLT4 and inhibits human colorectal cancer metastasis. Oncotarget. 2015 Feb;6(5):3225–39.

53. Tsai T-H, Yang C-C, Kou T-C, Yang C-E, Dai J-Z, Chen C-L, et al. Overexpression of GLUT3 promotes metastasis of triple-negative breast cancer by modulating the inflammatory tumor microenvironment. J Cell Physiol. 2021 Jun;236(6):4669–80.

54. Foran E, McWilliam P, Kelleher D, Croke DT, Long A. The leukocyte protein L-plastin induces proliferation, invasion and loss of E-cadherin expression in colon cancer cells. Int J cancer. 2006 Apr;118(8):2098–104.

55. Tiedemann K, Sadvakassova G, Mikolajewicz N, Juhas M, Sabirova Z, Tabariès S, et al. Exosomal Release of L-Plastin by Breast Cancer Cells Facilitates Metastatic Bone Osteolysis. Transl Oncol. 2019 Mar;12(3):462–74.

56. Yu Y, Wang W, Lu W, Chen W, Shang A. Inhibin β-A (INHBA) induces epithelial-mesenchymal transition and accelerates the motility of breast cancer cells by activating the TGF-β signaling pathway. Bioengineered. 2021 Dec;12(1):4681–96.

57. Bhat R, Thangavel H, Abdulkareem NM, Vasaikar S, De Angelis C, Bae L, et al. NPY1R exerts inhibitory action on estradiol-stimulated growth and predicts endocrine sensitivity and better survival in ER-positive breast cancer. Sci Rep. 2022 Feb;12(1):1972.

58. Cheng R, Qi L, Kong X, Wang Z, Fang Y, Wang J. Identification of the Significant

Genes Regulated by Estrogen Receptor in Estrogen Receptor-Positive Breast Cancer and Their Expression Pattern Changes When Tamoxifen or Fulvestrant Resistance Occurs. Front Genet. 2020;11:538734.

Table 1: Triads in ETR cell line model systems

Cell lines	mRNAs (UP)	lncRNA (UP)	miRNAs (DOWN)
TAMR AND LTEDI	CXCR4 SLC2A3 LCP1 RGS16 FLT4	LINC01588 LINC01825 MIR503HG AC027088.1	miR-182 miR-145 miR-424 miR-195 miR-497 miR-219-5p
	mRNAs (DOWN)	**lncRNAs (DOWN)**	**miRNAs (UP)**
	INHBA CALCR NPY1R DMD	AL645608.1 HAGLR AC027801.1 PPP3CB-AS1 AL137145.2 AC011450.1	miR-192 miR-215 miR-9 miR-140-5p miR-214 miR-31 miR-490-3p
	mRNAs (UP)	**lncRNAs (DOWN)**	**miRNAs (UP)**
	LCP1 CXCR4 RGS16 B3GNT3	AC079354.1 AL645608.1 AL137145.2	miR-205 miR-192 miR-215 miR-182 miR-214
	mRNAs (DOWN)	**lncRNAs (UP)**	**miRNAs (UP)**
	INHBA PPAP2B FHL1 CD36 IFIT1 DMD CXCL14	PCOTH AL162595.1	miR-9 miR-203
	mRNAs (UP)	**lncRNAs (UP)**	**miRNAs (UP)**
JOE	ADORA2B LCP1 RGS16	AC009951.1 MIR17HG SNHG14	miR-205 miR-182 miR-203

Table 2: miRNA expression levels in the plasma sample

Cell lines	miRNAs (UP)	miRNA expression in plasma		References
		Normal	Patient	
T A M R A N D L T E D I	miR-205	√	√	(44)
	miR-182	√	√	(45)
	miR-203	√	√	
	miR-214	√	√	
	miR-31	√	√	(45)
	miR-490-3p	√	√	(45)
	miR-9	√	√	(45)
	miR-140-5p	√	-	
	miR-192	√	-	
	miR-215	√	-	
	miRNAs (DOWN)			
	miR-182	√	√	(37,46)
	miR-145	√	-	
	miR-424	√	-	
	miR-195	√	-	
	miR-497	√	-	
	miR-219-5p	√	-	
	miRNAs (UP)			
J O E	miR-205	√	-	
	miR-182	√	-	
	miR-203	√	-	

Reference: Wang *et al*., 2012 (47)

Table 3: The seven upregulated genes and their role in cancers

Gene	Known function in cancer(s)	References
CXCR4	Highly expressed in tamoxifen-resistant cells and CXCR4 knockdown re-sensitized cells to endocrine therapy	(30,48)
B3GNT3	Downregulation of B3GNT3 was shown to enhance anti-neoplastic activity of cytotoxic T cells	(31,49,50)
FLT4	Associated with tumor progression in cervical cancer and highly expressed in tamoxifen resistance cells in breast cancer	(51,52)
SLC2A3	In triple-negative breast cancer, overexpression of SLC2A3 enhanced metastasis	(53)
LCP1	An oncogenic modulator that promotes tumor progression in colon cancer and its exosomal release in breast cancer improves osteolysis	(54,55)
INHBA	INHBA is highly expressed in breast tissue compared to normal and helps in cell motility, migration, and invasion	(56)
NPY1R	Overexpression of NPY1R positively correlated with breast cancer progression	(57)
CALCR	Highly expressed in tamoxifen-resistant cells	(58)

FIGURE LEGENDS

Figure 1: Effect of 4-OHT in LTEDI cells. **(A)** Cell proliferation assay of MCF7 and LTEDI cells in the presence and absence of 4-OHT measured by CCK8 kit. *$p \leq 0.05$, **$p \leq 0.01$, and ***$p \leq 0.001$. **(B)** Colony Forming Assay of TAMR and LTEDI cells in the presence and absence of 4-OHT.

Figure 2: Workflow for constructing the RNA regulatory networks.

Figure 3: Transcriptome analysis of cell lines. (A) DEGs were found in TAMR, LTEDI, and JOE cell lines compared to MCF7 cells. Numbers indicate DEGs in respective cell lines. **(B)** Upregulated and downregulated DEGs in TAMR and LTEDI cell lines. **(C)** Bubble plot showing **s**ignificant pathways found in TAMR and LTEDI cell lines. The X-axis shows a number of genes in each pathway and Y-axis indicates the name of the pathway. The pathways and bubble size key is indicated on the right-hand side. **(D)** Venn diagram of the intersection of TAMR, LTEDI, and JOE DEGs. **(E)** Heatmap of 170 genes with green (downregulated) to red (upregulated) DEGs. A Chord diagram showing pathway curated genes (7 upregulated and 14 downregulated) in multiple pathways. Significant pathways are shown on the right, and the fold change of core genes is shown on the left. Left–right connections indicate gene membership in a pathway's leading-edge subset.

Figure 4: Estimation of DELs in three cell lines. (A) Differentially expressed long noncoding RNAs, (**B**) upregulated and downregulated long noncoding RNAs in TAMR, LTEDI, and JOE cell lines.

Figure 5: Construction of co-expression network. CNC network of TAMR and LTEDI cells using **(A)** Upregulated DELs and **(B)** Downregulated DELs co-expressed with 7 upregulated DEGs. **(B)** 14 downregulated DEGs correlated with 93 common DELs. **(C)** Upregulated and **(D)** Downregulated DEL networks in JOE cells. The ellipse represents DEGs, and diamonds represent DELs.

Figure 6: Construction of competing endogenous RNA (ceRNA) network using (A) mRNAs, lncRNAs (up), and miRNAs (down) in the TAMR/LTEDI network, **(B**) mRNAs, lncRNAs (down), and miRNAs (up) in the TAMR/LTEDI network, **(C)** Network of upregulated mRNAs, miRNAs, and downregulated lncRNAs in TAMR and LTEDI cell lines, **(D)** mRNAs (down), lincRNAs, and miRNAs (up) in TAMR and LTEDI cell lines and **(E)** ceRNA network constructed in JOE cells using upregulated pairs. The ellipse represents mRNAs, the diamond represents lncRNAs, "V" shape represents miRNAs. For the prediction of miRNA levels- Red represents upregulated, and green/blue represents downregulated miRNAs.

Figure 7: Validation of predicted miRNA expression. Publicly available miRNA expression data from GSE56411 from 3 different TAMR clones was compared with the prediction of miRNA regulation from this study. Validated miRNAs are marked circled in Figure 7.

Figure 8: Gene expression in breast cancer patient samples. (A) Supervised hierarchical clustering analysis of 1084 samples (all subtypes) using mRNA expression z-scores relative to all samples (log RNA Seq. V2 RSEM) for a 73-gene ETR cassette. **(B)** Expression of upregulated genes of normal and primary tumor TCGA samples using the UALCAN database analysis (**** $p<0.01$). **(C)** Kaplan-Meier survival curves of select genes are shown.

Supplementary figure 1: RNA-seq analysis of Fulvestrant data (GSE118713). (A) Venn diagram of the common genes between TAMR, LTEDI, JOE, and GSE118713. (B) Expression patterns of 85 genes are shown by heatmap. Changes in expression levels are displayed from green (downregulated) to red (upregulated), as shown in the color gradient on the right side.

Supplementary figure 2: CNC network of individual TAMR and LTEDI cells. (A and B) upregulated DEGs:DELs and (C and D) downregulated DEGs:DELs in TAMR and LTEDI cells respectively.

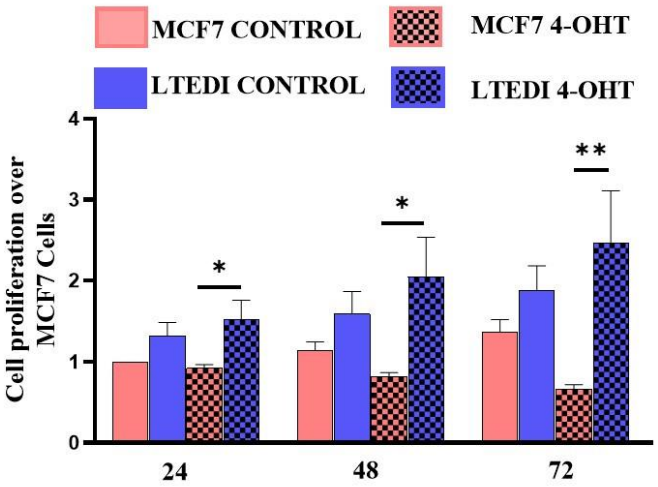
Supplementary figure 3: Construction of competing endogenous RNA (ceRNA) network. The upregulated ceRNA network of TAMR-LTEDI cells (A) upregulated, (B) downregulated. (C) ceRNA network of upregulated DEGs and downregulated DELs from TAMR and LTEDI cells. (D) ceRNA network of downregulated DEGs and upregulated DELs from TAMR and LTEDI cells. The ellipse represents mRNAs, the diamond represents lncRNAs, "V" shape represents miRNAs.

Supplementary figure 4: Gene expression in breast cancer patient samples from the publicly available dataset. **(A)** Supervised hierarchical clustering analysis of 1084 samples (all subtypes) using mRNA expression of downregulated genes (log RNA Seq. V2 RSEM). **(B)** Expression of upregulated genes of normal and primary tumor TCGA samples using the UALCAN database analysis (**** $P<0.01$). **(C)** Kaplan-Meier survival curves of select genes.

Supplementary file 1: List of DEGs and DELs identified in TAMR and LTEDI cells (adjusted $p \leq 0.05$)

Figure 1

A


MCF7 CONTROL
MCF7 4-OHT
LTEDI CONTROL
LTEDI 4-OHT
Cell proliferation over MCF7 Cells
0
1
2
3
4
*
*
**
24
48
72

B

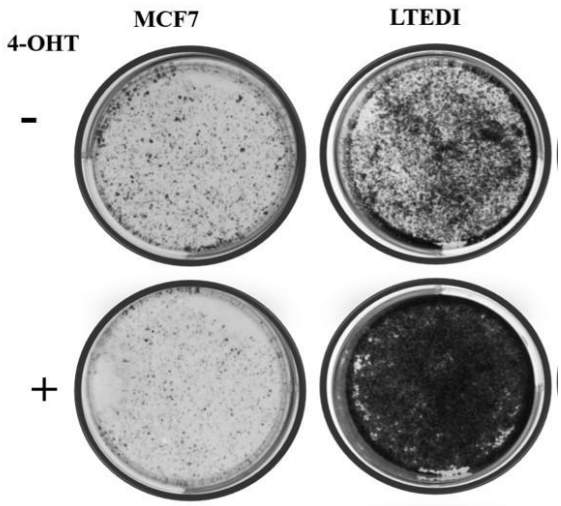

4-OHT
MCF7
LTEDI
-
+

Figure 2

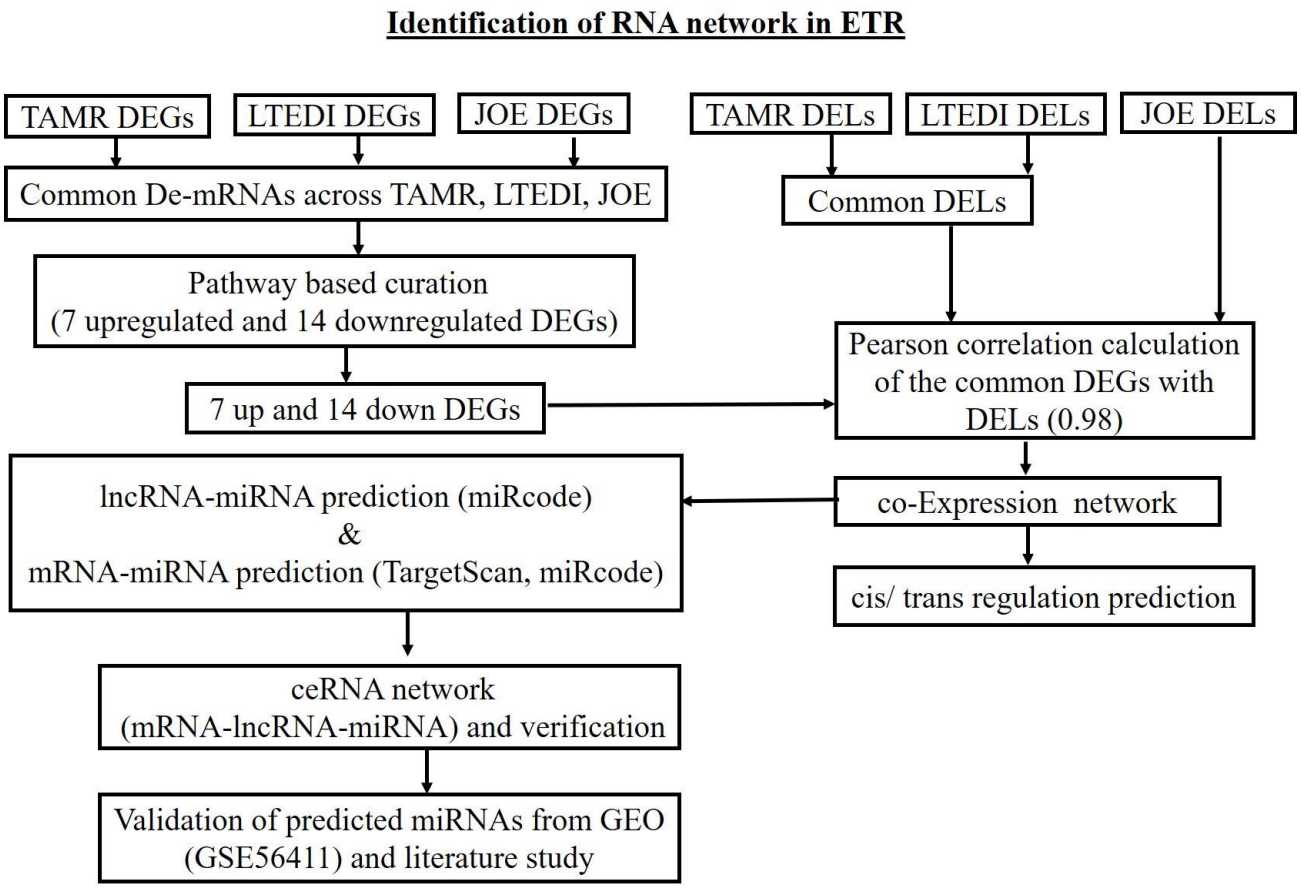

Identification of RNA network in ETR
TAMR DEGs
LTEDI DEGs
JOE DEGs
TAMR DELs
LTEDI DELs
JOE DELs
Common De-mRNAs across TAMR, LTEDI, JOE
Common DELs
Pathway based curation
(7 upregulated and 14 downregulated DEGs)
Pearson correlation calculation of the common DEGs with DELs (0.98)
7 up and 14 down DEGs
lncRNA-miRNA prediction (miRcode)
&
mRNA-miRNA prediction (TargetScan, miRcode)
co-Expression network
cis/ trans regulation prediction
ceRNA network
(mRNA-lncRNA-miRNA) and verification
Validation of predicted miRNAs from GEO
(GSE56411) and literature study

Figure 3

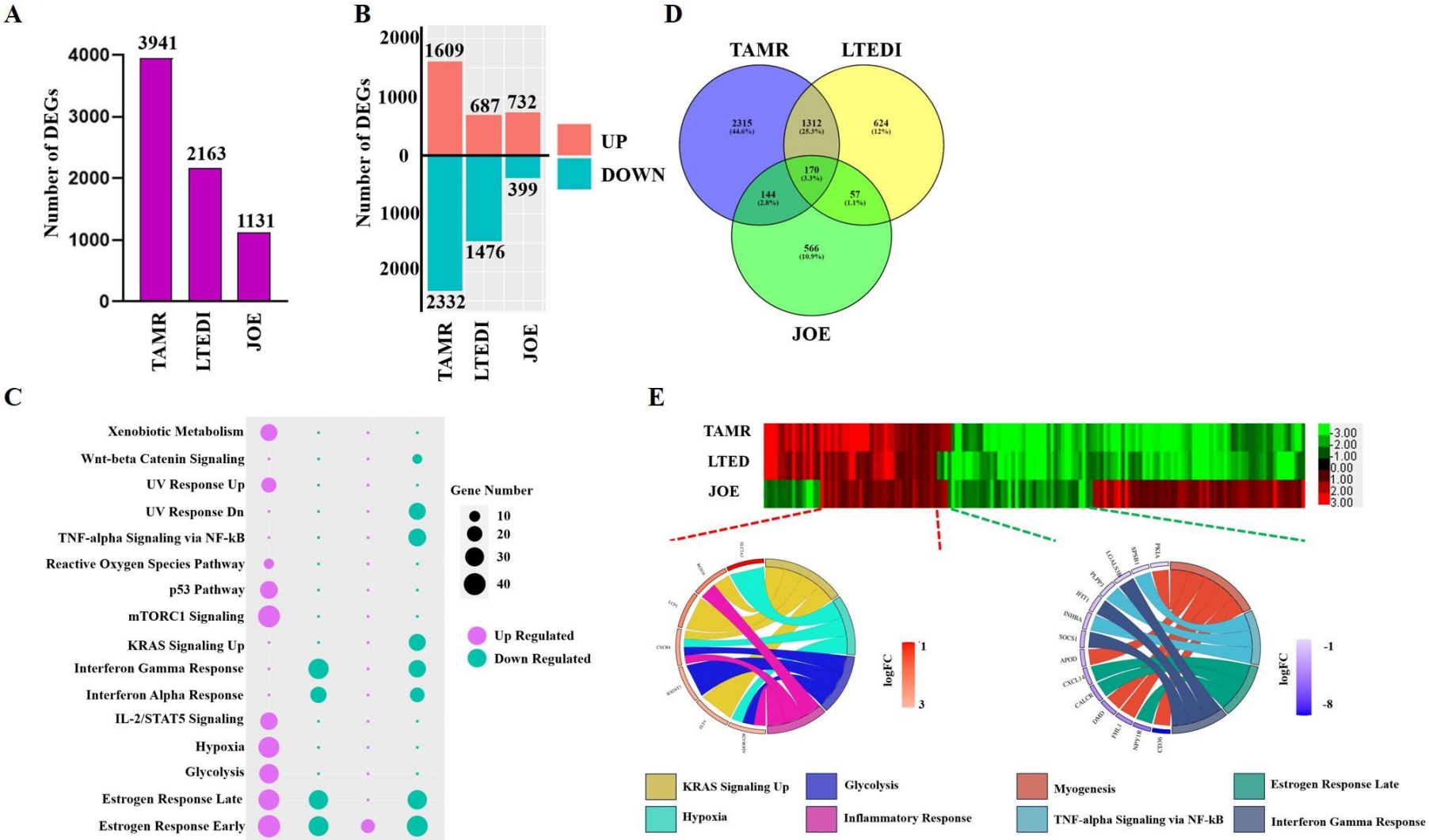

A
Number of DEGs
4000
3000
2000
1000
0
3941
2163
1131
TAMR
LTEDI
JOE
B
Number of DEGs
2000
1000
0
1000
2000
1609
687
732
399
1476
2332
UP
DOWN
TAMR
LTEDI
JOE
D
TAMR
LTEDI
JOE
2315 (44.6%)
1312 (25.3%)
624 (12%)
170 (3.3%)
144 (2.8%)
57 (1.1%)
566 (10.9%)
C
Xenobiotic Metabolism
Wnt-beta Catenin Signaling
UV Response Up
UV Response Dn
TNF-alpha Signaling via NF-kB
Reactive Oxygen Species Pathway
p53 Pathway
mTORC1 Signaling
KRAS Signaling Up
Interferon Gamma Response
Interferon Alpha Response
IL-2/STAT5 Signaling
Hypoxia
Glycolysis
Estrogen Response Late
Estrogen Response Early
Epithelial Mesenchymal Transition
Cholesterol Homeostasis
Up
Down
Up
Down
TAMR
LTEDI
Gene Number
10
20
30
40
Up Regulated
Down Regulated
E
TAMR
LTED
JOE
3.00
2.00
1.00
0.00
1.00
2.00
3.00
logFC
1
3
logFC
-1
-8
KRAS Signaling Up
Hypoxia
Glycolysis
Inflammatory Response
Myogenesis
TNF-alpha Signaling via NF-kB
Estrogen Response Late
Interferon Gamma Response

Figure 4

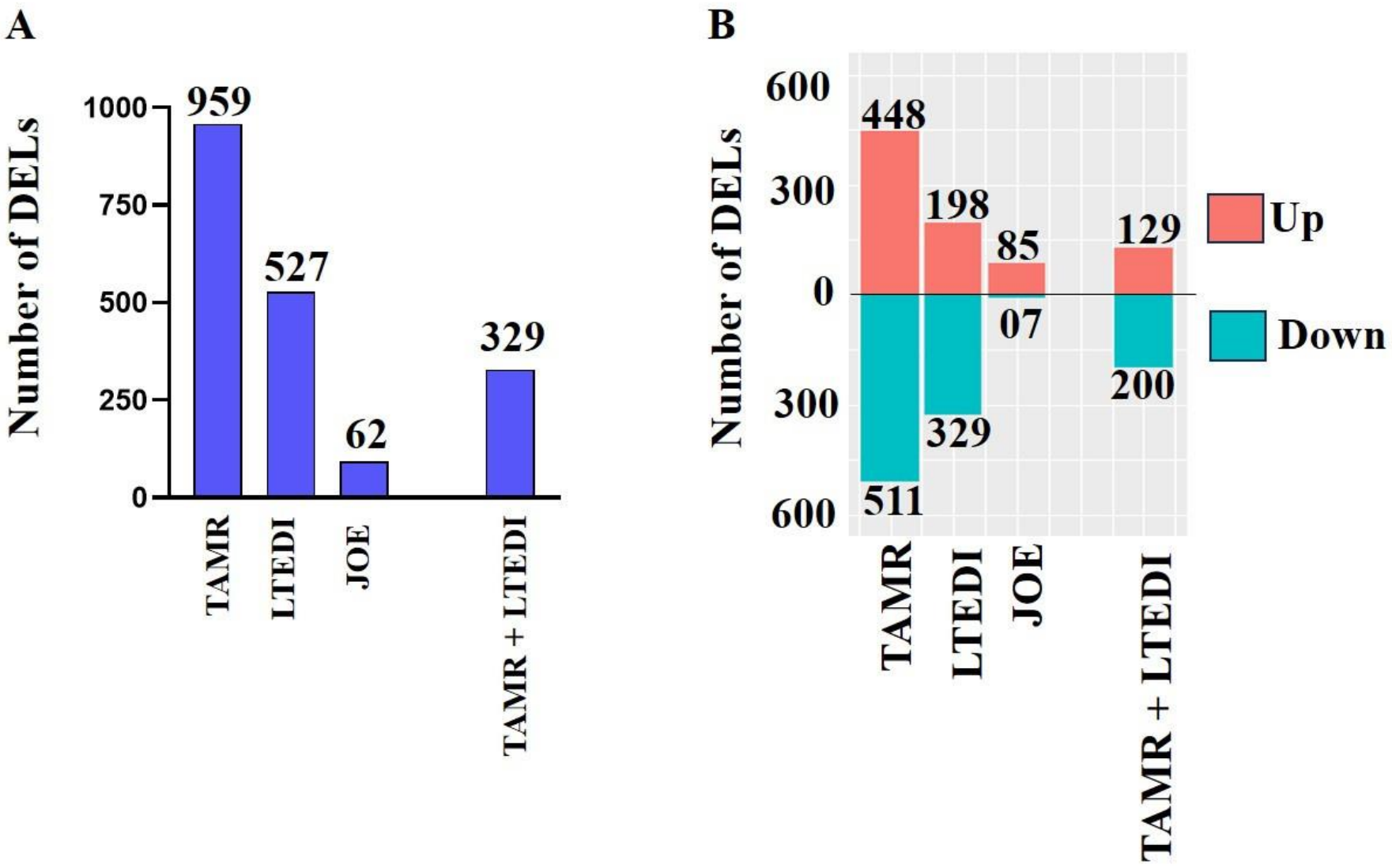

A
Number of DELs
1000
750
500
250
0
959
527
62
329
TAMR
LTEDI
JOE
TAMR + LTEDI
B
Number of DELs
600
300
0
300
600
448
198
85
129
07
329
511
200
Up
Down
TAMR
LTEDI
JOE
TAMR + LTEDI

Figure 5

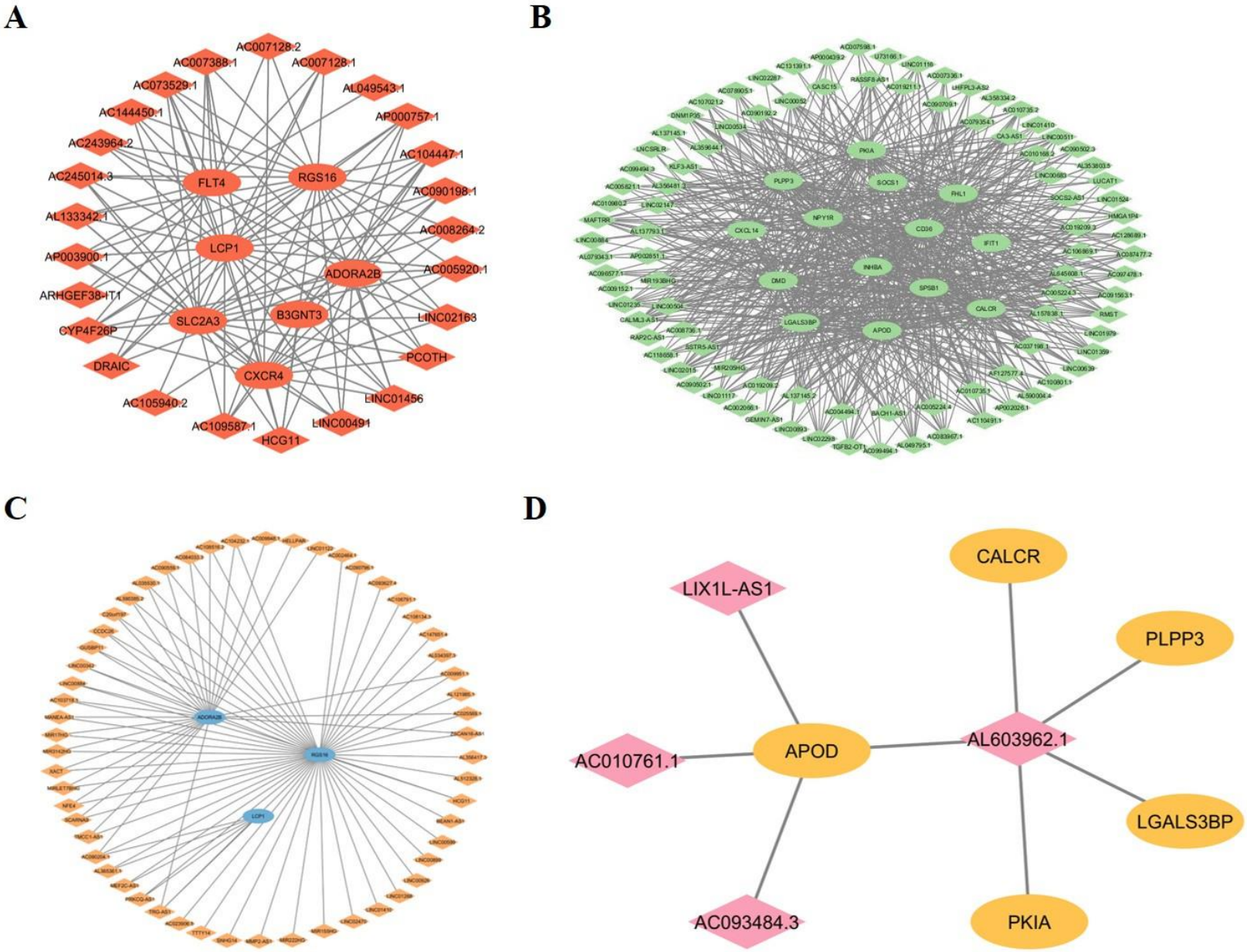

A
AC007128.2
AC007388.1
AC007128.1
AC073529.1
AL049543.1
AC144450.1
AP000757.1
AC243964.2
AC104447.1
AC245014.3
FLT4
RGS16
AC090198.1
AL133342.1
AC008264.2
LCP1
AP003900.1
ADORA2B
AC005920.1
ARHGEF38-IT1
SLC2A3
B3GNT3
LINC02168
CYP4F26P
PCOTH
DRAIC
CXCR4
LINC01456
AC105940.2
LINC00491
AC109587.1
HCG11
B
PKIA
PLPP3
SOCS1
FHL1
NPY1R
CD36
CXCL14
IFIT1
INHBA
DMD
SPSB1
CALCR
LGALS3BP
APOD
C
D
CALCR
LIX1L-AS1
PLPP3
AC010761.1
APOD
AL603962.1
LGALS3BP
AC093484.3
PKIA

Figure 6

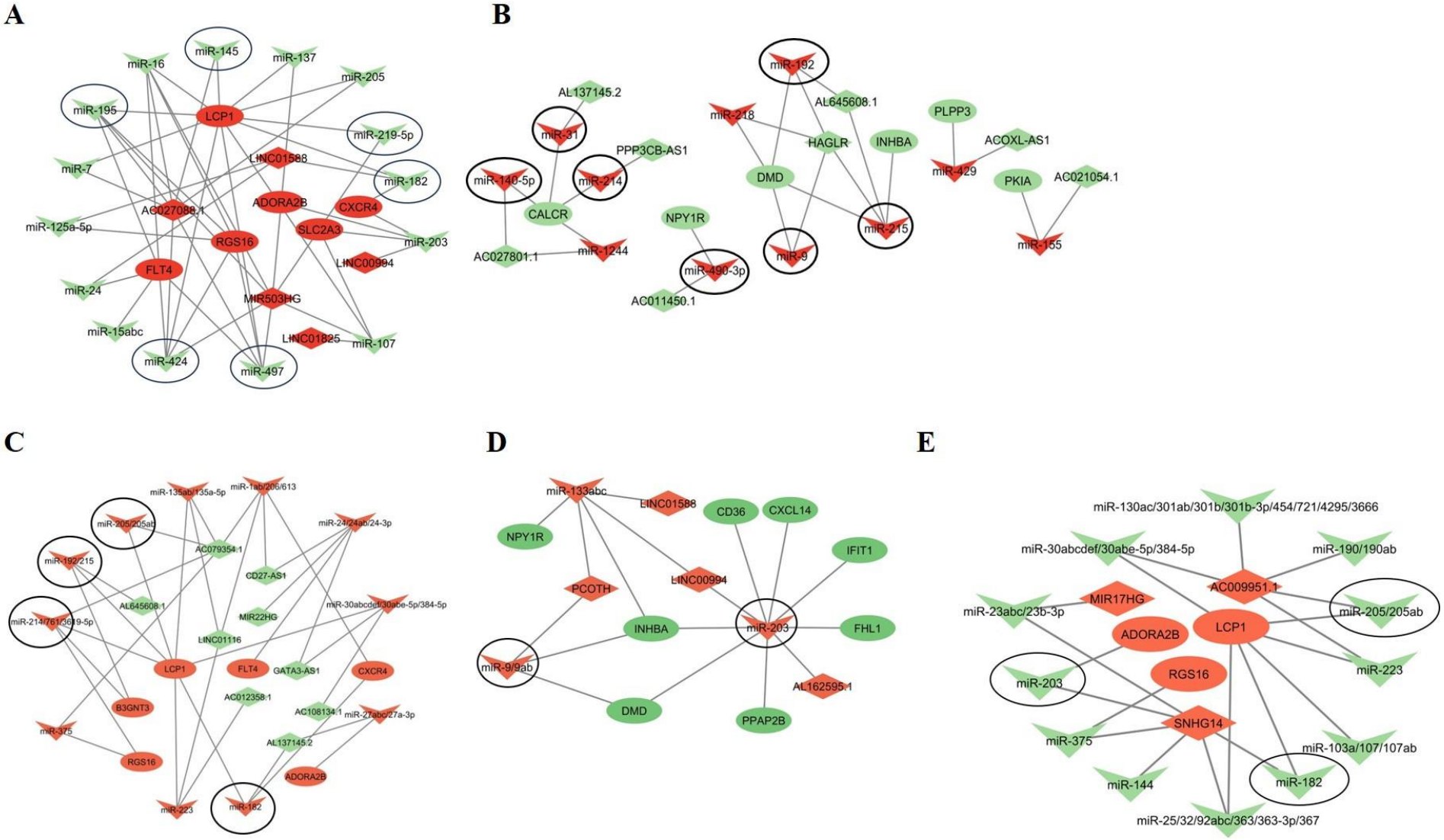

A
miR-145
miR-16
miR-137
miR-205
miR-195
LCP1
miR-219-5p
miR-7
LINC01588
miR-182
AC027088.1
ADORA2B
CXCR4
miR-125a-5p
SLC2A3
RGS16
miR-203
LINC00994
FLT4
miR-24
MIR503HG
miR-15abc
LINC01825
miR-107
miR-424
miR-497
B
miR-192
AL137145.2
AL645608.1
miR-218
PLPP3
miR-31
HAGLR
INHBA
ACOXL-AS1
PPP3CB-AS1
miR-429
miR-140-5p
miR-214
DMD
PKIA
AC021054.1
CALCR
NPY1R
miR-215
AC027801.1
miR-1244
miR-9
miR-155
miR-490-3p
AC011450.1
C
miR-135ab/135a-5p
miR-1ab/206/613
miR-205/205ab
miR-24/24ab/24-3p
AC079354.1
miR-192/215
CD27-AS1
AL645608.1
miR-30abcdef/30abe-5p/384-5p
miR-214/761/3619-5p
MIR22HG
LINC01116
LCP1
FLT4
GATA3-AS1
CXCR4
AC012358.1
B3GNT3
AC108134.1
miR-27abc/27a-3p
miR-375
AL137145.2
RGS16
ADORA2B
miR-223
miR-182
D
miR-133abc
LINC01588
CD36
CXCL14
NPY1R
IFIT1
LINC00994
PCOTH
INHBA
miR-203
FHL1
miR-9/9ab
AL162595.1
DMD
PPAP2B
E
miR-130ac/301ab/301b/301b-3p/454/721/4295/3666
miR-30abcdef/30abe-5p/384-5p
miR-190/190ab
AC009951.1
MIR17HG
miR-23abc/23b-3p
miR-205/205ab
ADORA2B
LCP1
miR-223
miR-203
RGS16
SNHG14
miR-375
miR-103a/107/107ab
miR-144
miR-182
miR-25/32/92abc/363/363-3p/367

Figure 7

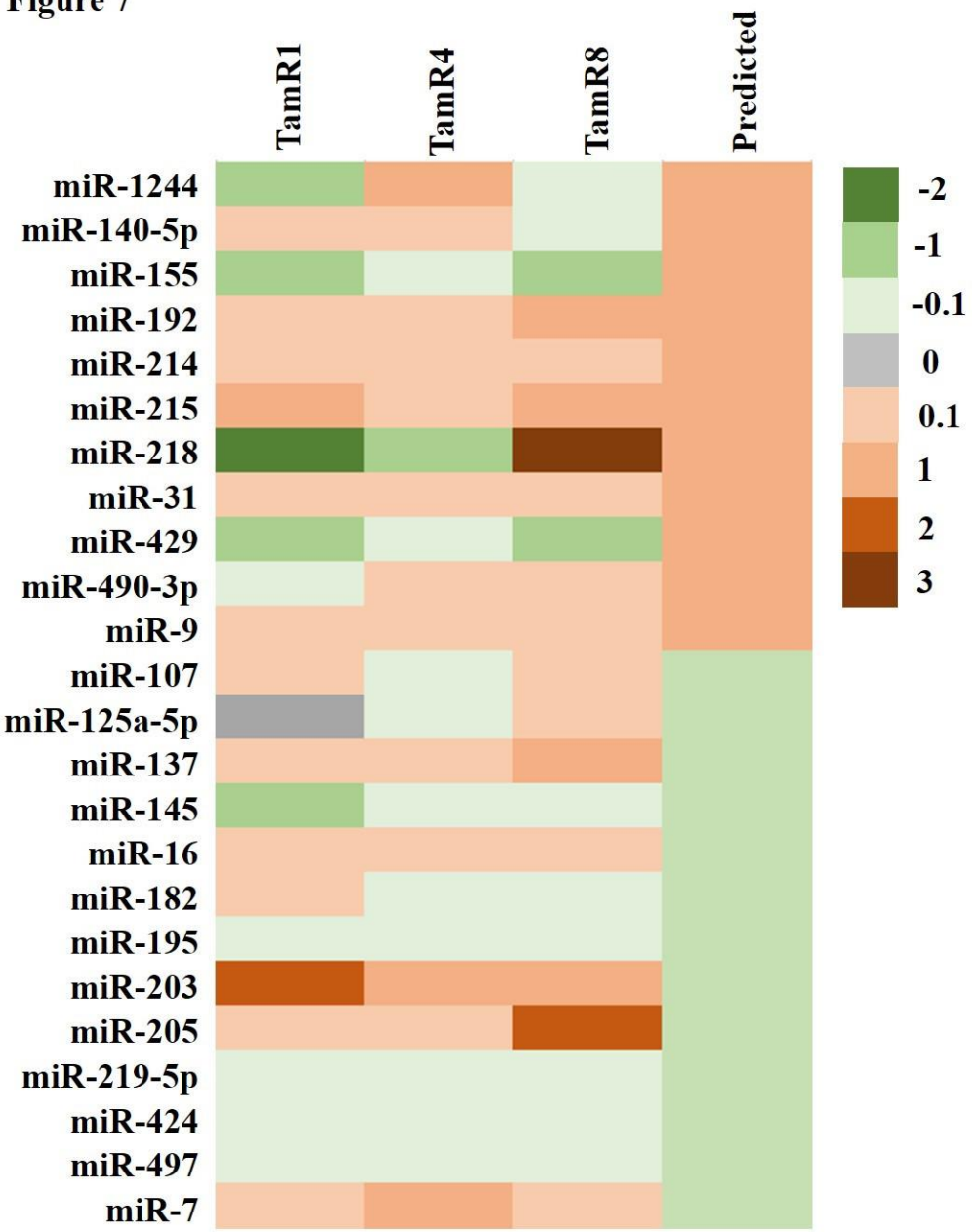
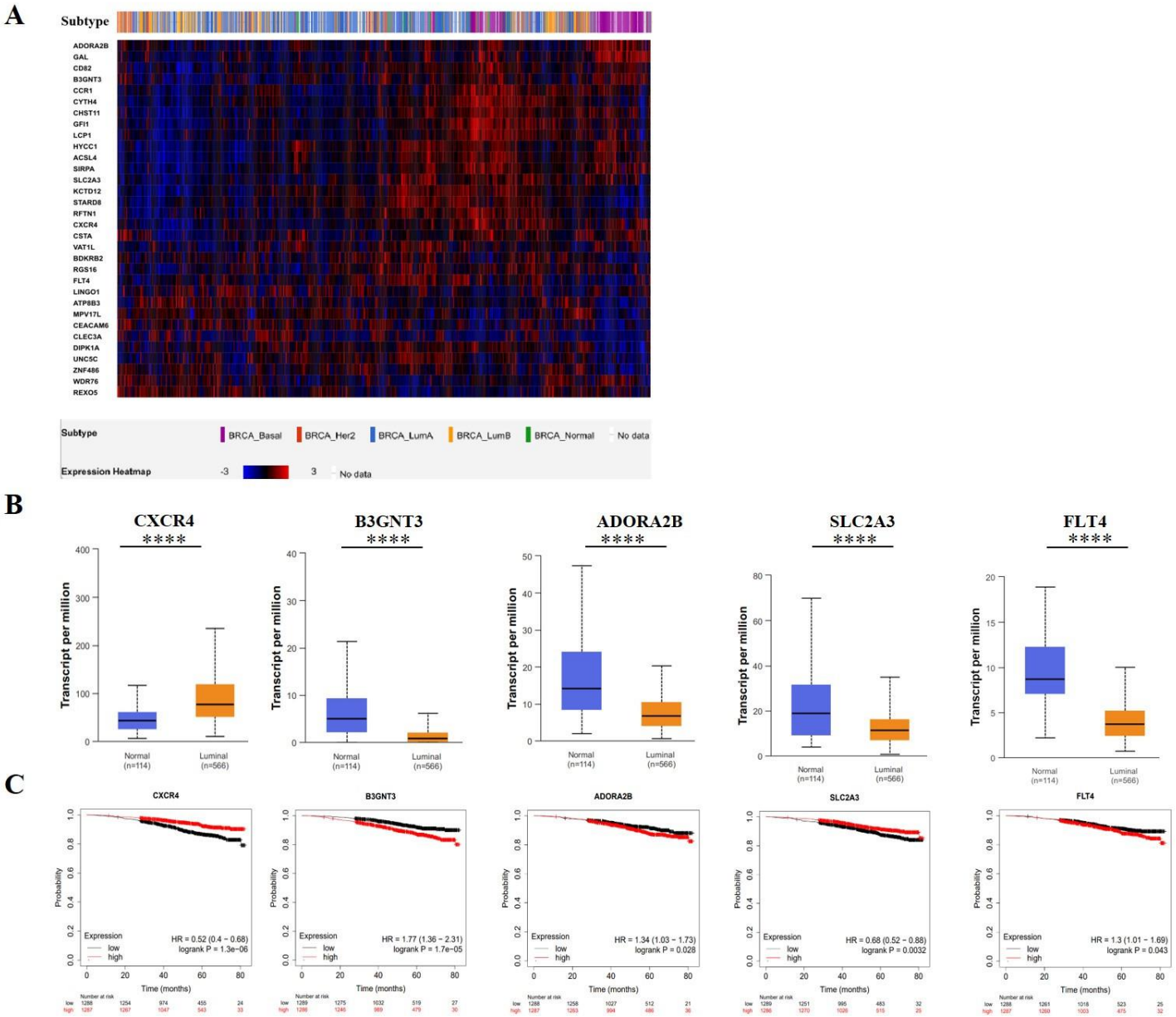

TamR1
TamR4
TamR8
Predicted
miR-1244
miR-140-5p
miR-155
miR-192
miR-214
miR-215
miR-218
miR-31
miR-429
miR-490-3p
miR-9
miR-107
miR-125a-5p
miR-137
miR-145
miR-16
miR-182
miR-195
miR-203
miR-205
miR-219-5p
miR-424
miR-497
miR-7
-2
-1
-0.1
0
0.1
1
2
3

Figure 8


A
Subtype
ADORA2B
GAL
CD82
B3GNT3
CCR1
CYTH4
CHST11
GFI1
LCP1
HYCC1
ACSL4
SIRPA
SLC2A3
KCTD12
STARD8
RFTN1
CXCR4
CSTA
VAT1L
BDKRB2
RGS16
FLT4
LINGO1
ATP8B3
MPV17L
CEACAM6
CLEC3A
DIPK1A
UNC5C
ZNF486
WDR76
REXO5
Subtype
BRCA_Basal
BRCA_Her2
BRCA_LumA
BRCA_LumB
BRCA_Normal
No data
Expression Heatmap
-3
3
No data
B
CXCR4
****
B3GNT3
****
ADORA2B
****
SLC2A3
****
FLT4
****
Transcript per million
Normal (n=114)
Luminal (n=566)
C
CXCR4
B3GNT3
ADORA2B
SLC2A3
FLT4
Probability
Expression
low
high
HR = 0.52 (0.4 − 0.68)
logrank P = 1.3e−06
HR = 1.77 (1.36 − 2.31)
logrank P = 1.7e−05
HR = 1.34 (1.03 − 1.73)
logrank P = 0.028
HR = 0.68 (0.52 − 0.88)
logrank P = 0.0032
HR = 1.3 (1.01 − 1.69)
logrank P = 0.043
Time (months)
Number at risk